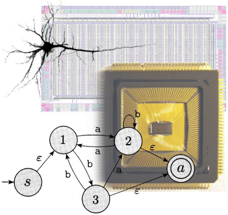
UCI

Scalable Neuromorphic Learning Machines

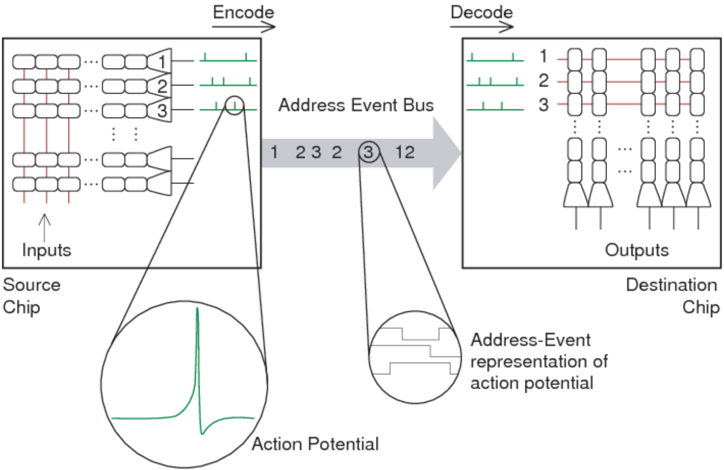
Emre Neftci

Department of Cognitive Sciences, UC Irvine,
Department of Computer Science, UC Irvine,

September 11, 2017



Neuromorphic Computing Can Enable Low-power, Massively Parallel Computing



- Only spikes are communicated & routed between neurons (weights, internal states are local)
- To use this architecture for practical workloads, we need algorithms that operate on local information

System-level Approach

Objective Function: Target spike train s^*

$$\mathcal{L}(\mathbf{s}^*, \mathbf{s}, \boldsymbol{w})$$

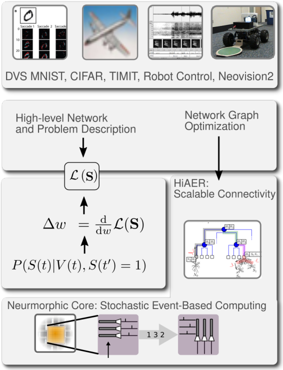
Neuron Model:
Probability of spike given input spike train $\mathbf{s}$

$$P(s_i = 1 | \mathbf{s}) = \rho(u_i)$$

$$u_i(t) = \sum_j w_{ij}\, \epsilon * s_j(t) + \eta * s_i(t)$$

"Activation function" ρ_i can be derived or estimated. Kernels η and ϵ reflect neural and synaptic dynamics.

Gerstner and Kistler,, 2002

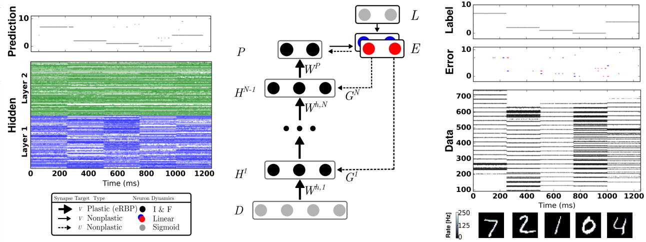


Learning Rule:

$$\Delta w \propto \frac{\partial}{\partial w_{ij}}\mathcal{L} = \frac{\partial \mathcal{L}}{\partial s_i}\rho'(u_i)\frac{\partial u_i}{\partial w_{ij}}$$

Spiking neural networks can be viewed as (deep) Binary Neural Networks

Event-Driven Random Backpropagation (eRBP) for Deep Supervised Learning



function ERBP

  for $k \in \{\text{presynaptic spike addresses } \mathbf{S}^{pre}\}$ **do**

    if $b_{min} < I < b_{max}$ **then** $w_k \leftarrow w_k + T$,

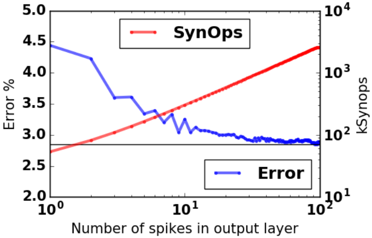
    end if

  end for

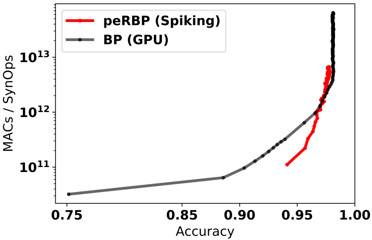
end function

Dataset	eRBP$_\times$	BP (100)
MNIST 784-200-200-10	**2.29** %	1.80 %
MNIST 784-500-500-10	**2.02** %	1.90 %

- **Energy Efficiency During Inference:** First output spike is >95% accurate



- **Energy Efficiency During Training:** SynOp-MAC parity



Neftci, Augustine, Paul, and Detorakis, *Frontiers in Neuroscience*, 2017

Digital hardware architecture for neuromorphic learning machines (HiAER NSAT)

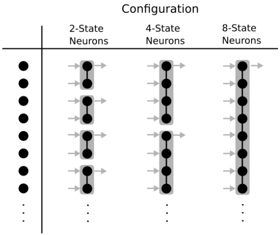
(a) Neural and Synaptic Array Transeiver

Neuron Dynamics

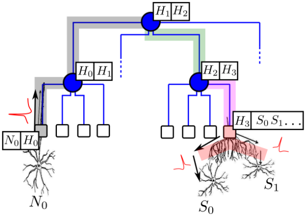
$$
\begin{aligned}
\mathbf{x}[t+1] &= \mathbf{A}\mathbf{x}[t] && \textbf{(Leak \& Coupling)}\\
&\quad + \boldsymbol{\Xi}[t] \otimes \mathbf{W}[t]\mathbf{S}[t] && \textbf{(Synaptic inputs)}\\
&\quad + \boldsymbol{\eta}[t] && \textbf{(Noise)}\\
x_0[t+1] &\geq \theta_0, s_i[t+1] \leftarrow 1 && \textbf{(Spiking Output)}\\
\mathbf{x}[t+1] &\geq \boldsymbol{\theta}, \mathbf{x}[t+1] \leftarrow \mathbf{X_r} && \textbf{(Thresholds \& Reset)}
\end{aligned}
$$

Synaptic Plasticity Dynamics

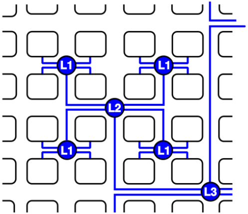
$$
\begin{aligned}
e_k &= x_m[t]\,(K[t-t_k] + K[t_k - t_{last}]) && \textbf{(Eligibilty)}\\
w_k[t+1] &= w_k[t] + s_k[t+1]e_k && \textbf{(Weight update)}
\end{aligned}
$$



(b) Connectivity Model

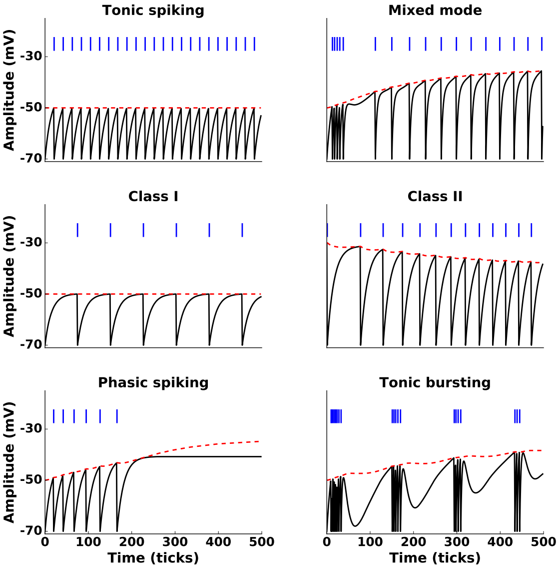


(c) HiAER Tree



Detorakis, Augustine, Paul, Pedroni, Sheik, Cauwenberghs, and Neftci (in preparation)

NSAT Neural Dynamics Flexibility



Detorakis, Augustine, Paul, Pedroni, Sheik, Cauwenberghs, and Neftci (in preparation)

NSAT Learning Dynamics Flexibility

$$w_k[t+1] = w_k[t] + s_k[t+1]e_k \qquad \text{(Weight update)}$$

$$e_k = x_m \underbrace{\left(K[t-t_k] + K[t_k - t_{last}]\right)}_{STDP} \qquad \text{(Eligibilty)}$$

$$x_m = \sum_i \gamma_i x_i \qquad \text{(Modulation)}$$

Detorakis, Augustine, Paul, Pedroni, Sheik, Cauwenberghs, and Neftci (in preparation)

NSAT Learning Dynamics Flexibility

$$w_k[t+1] = w_k[t] + s_k[t+1]e_k \qquad \text{(Weight update)}$$

$$e_k = x_m \underbrace{\left(K[t-t_k] + K[t_k - t_{last}]\right)}_{STDP} \qquad \text{(Eligibilty)}$$

$$x_m = \sum_i \gamma_i x_i \qquad \text{(Modulation)}$$

Detorakis, Augustine, Paul, Pedroni, Sheik, Cauwenberghs, and Neftci (in preparation)

STDP weight updates on pre-synaptic spikes only, using only forward lookup access of the synaptic connectivity table

Pedroni et al.,, 2016

"Plasticity involves as a third factor a local dendritic potential, besides pre- and postsynaptic firing times"

Urbanczik and Senn, *Neuron*, 2014

Clopath, Büsing, Vasilaki, and Gerstner, *Nature Neuroscience*, 2010

- **Reinforcement Learning**

$$\Delta w_{ij} = \eta\, rSTDP_{ij}$$

Florian, *Neural Computation*, 2007

- **Unsupervised Representation Learning**

$$\Delta w_{ij} = \eta\, g(t) STDP_{ij}$$

Neftci, Das, Pedroni, Kreutz-Delgado, and Cauwenberghs, *Frontiers in Neuroscience*, 2014

- **Unsupervised Sequence Learning**

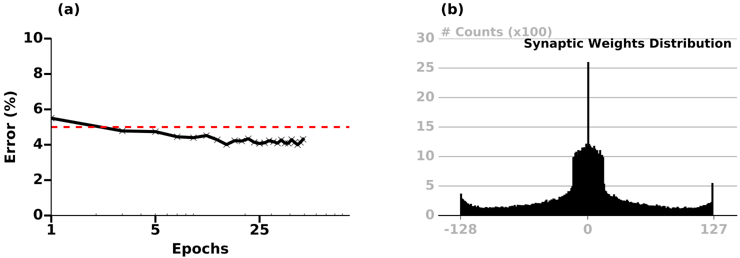
$$\Delta w_{ij} = \eta\, (\Theta(V) - \alpha(\nu_i - C))\, s_j$$

Sheik et al. 2016

- **Supervised Deep Learning (eRBP)**

$$\Delta w_{ij} = \eta\, Error_i \phi'(t) s_j$$

Neftci, Pedroni, Joshi, Al-Shedivat, and Cauwenberghs, *Frontiers in Neuroscience*, 2016



- Event-driven Random Backpropagation Rule
- MNIST 784-100-10
- 8 bit synaptic weights
- 16 bit neural states

Detorakis, Augustine, Paul, Pedroni, Sheik, Cauwenberghs, and Neftci (in preparation)

Acknowledgements

Collaborators:

Gert Cauwenberghs

(UCSD)

Georgios Detorakis

(UCI)

Somnath Paul

(Intel)

Charles Augustine

(Intel)

Sponsors:

- National Science Foundation Grant No. 1652159 “CAREER: Scalable Neuromorphic Learning Machines”, No. 1640081
- SRC-NRI Nanoelectronics Research Initiative
- Intel Corporation, Intel Strategic Research Alliance
- KIST Neuromorphic Research Consortium

C. Clopath, L. Büsing, E. Vasilaki, and W. Gerstner. “Connectivity reflects coding: a model of voltage-based STDP with homeostasis”. In: *Nature Neuroscience* 13.3 (2010), pp. 344–352.

R.V. Florian. “Reinforcement learning through modulation of spike-timing-dependent synaptic plasticity”. In: *Neural Computation* 19.6 (2007), pp. 1468–1502.

W. Gerstner and W. Kistler. *Spiking Neuron Models. Single Neurons, Populations, Plasticity*. Cambridge University Press, 2002.

E. Neftci, S. Das, B. Pedroni, K. Kreutz-Delgado, and G. Cauwenberghs. “Event-Driven Contrastive Divergence for Spiking Neuromorphic Systems”. In: *Frontiers in Neuroscience* 7.272 (Jan. 2014). ISSN: 1662-453X. DOI: 10.3389/fnins.2013.00272. URL: http://www.frontiersin.org/neuromorphic_engineering/10.3389/fnins.2013.00272/abstract.

Emre O Neftci, Bruno Umbria Pedroni, Siddharth Joshi, Maruan Al-Shedivat, and Gert Cauwenberghs. “Stochastic Synapses Enable Efficient Brain-Inspired Learning Machines”. In: *Frontiers in Neuroscience* 10.241 (2016). ISSN: 1662-453X. DOI: 10.3389/fnins.2016.00241. URL: http://www.frontiersin.org/neuromorphic_engineering/10.3389/fnins.2016.00241/abstract.

Emre Neftci, Charles Augustine, Somnath Paul, and Georgios Detorakis. “Event-driven Random Back-Propagation: Enabling Neuromorphic Deep Learning Machines”. In: *Frontiers in Neuroscience* (May 2017). Accepted. URL: `arXiv%20preprint%20arXiv:1612.05596`.

Bruno U Pedroni, Sadique Sheik, Siddharth Joshi, Georgios Detorakis, Somnath Paul, Charles Augustine, Emre Neftci, and Gert Cauwenberghs. “Forward Table-Based Presynaptic Event-Triggered Spike-Timing-Dependent Plasticity”. In: Oct. 2016. URL: `%7BIEEE%20Biomedical%20Circuits%20and%20Systems%20Conference%20(BioCAS),%20https://arxiv.org/abs/1607.03070%7D`.

Robert Urbanczik and Walter Senn. “Learning by the dendritic prediction of somatic spiking”. In: *Neuron* 81.3 (2014), pp. 521–528.